

修 士 論 文 の 和 文 要 旨

研究科・専攻	大学院 情報理工 学研究科 総合情報学 専攻 博士前期課程		
氏 名	小倉 秀隆	学籍番号	1030021
論 文 題 目	アクセント型識別を用いた単語音声認識		
<p>要 旨</p> <p>近年、音声認識機能を搭載した製品が増えている。しかし、これらの製品の音声認識機能を使用する際にストレスを感じないとは言い難い。従来の音声認識では音韻情報が用いられてきたが、韻律情報も併用することで認識性能が向上する可能性が報告されている。そこで本研究では、認識処理に韻律情報を取り入れ、音声認識率の向上を目指した。</p> <p>本研究では、認識対象を単語音声とし、音韻情報と韻律情報を別系列で処理し、最後にその2つの処理結果を1つにまとめ、認識を行う枠組を提案する。以下に提案法の概要を示す。</p> <p>音韻情報の特徴量をメル周波数ケプストラム係数(MFCC)とした。認識処理を行う前に、学習用単語から切り出した音素の MFCC を用いて、音素モデルを学習する。認識処理では、入力単語リストを参照し音素モデルを組み合わせ、入力される可能性のある全ての単語のモデルを作る。これらのモデル全てに対し入力単語音声の音韻尤度を求める。</p> <p>韻律情報には声の高さ、大きさ、時間長などいくつかあるが、提案法ではアクセント型を用いた。まず入力単語音声の基本周波数値を抽出する。次に、基本周波数パターンの生成過程のモデル(以下、F_0 モデルとする)のパラメータを変化させ、F_0 モデルで生成される基本周波数変化の形状を入力音声の基本周波数変化の形状に当てはめる。当てはめた F_0 モデルのパラメータを提案法におけるアクセント型特徴量とした。アクセント型は、単語のモーラ数によって型の数が決まる。そこで、学習段階では、学習用単語から抽出したアクセント型特徴量を用いて、モーラ数ごとに全てのアクセント型モデルを学習する。認識処理では、それぞれのモデルに対し入力単語音声の韻律尤度を求める。</p> <p>入力単語リストに載っている単語全てについて、音韻尤度と韻律尤度をそれぞれ求め、それらを線形結合し、単語の尤度を得る。その尤度が最大となるモデルに対応する単語を入力単語と判定する。</p> <p>提案法の評価は、まずアクセント型識別実験を行い、つぎに韻律情報を用いた単語音声認識を行った。</p> <p>アクセント型識別実験では、従来法に比べ、クリーン音声では 20 話者の平均の識別率が向上した。種々の環境雑音を様々な信号対雑音比で重畳し評価した場合では、識別率は多くの条件で改善が見られたが、従来法とほとんど変わらないものも見られた。</p> <p>韻律情報を用いた単語音声認識実験では、従来法に比べほとんどの条件で認識率が向上した。</p>			